PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07K 14/81, A61K 38/55

A1

(11) International Publication Number: WO 00/64945

(43) International Publication Date: 2 November 2000 (02.11.00)

(21) International Application Number: PCT/GB00/01571

(22) International Filing Date: 20 April 2000 (20.04.00)

(30) Priority Data:
9909133.2 22 April 1999 (22.04.99) GB

(71) Applicant *(for all designated States except US)*: BABRAHAM INSTITUTE [GB/GB]; Babraham Hall, Babraham, Cambridge CB2 4AT (GB).

(72) Inventors; and

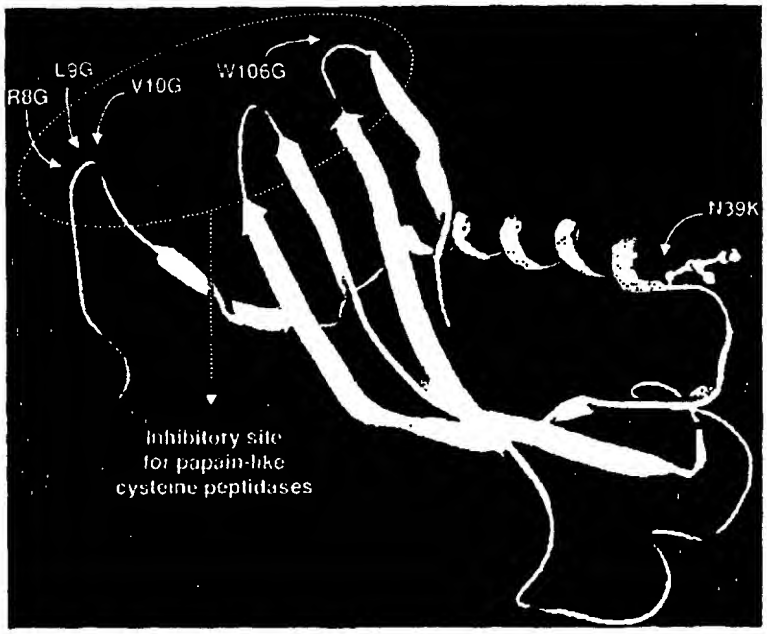
(75) Inventors/Applicants *(for US only)*: ABRAHAMSON, Magnus [SE/SE]; Aspinge 1801, S–241 94 Eslov (SE). BARRETT, Alan, John [GB/GB]; 8 Stansgate Avenue, Cambridge CB1 2QZ (GB).

(74) Agent: DAVIES, Jonathan, Mark; Reddie & Grose, 16 Theobalds Road, London WC1X 8PL (GB).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: INHIBITION OF ENDOPEPTIDASE BY CYSTATINS



(57) Abstract

The present invention relates to inhibition of legumain or legumain related endopeptidase by cystatin. The inhibitory effect on legumain is by virtue of activity of a "second site" on cystatin, not previously known. The "first site" on cystatin is known to bind papain and papain–related peptidase. Papain does not bind the "second site" on cystatin.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INHIBITION OF ENDOPEPTIDASE BY CYSTATINS

Introduction

The present invention relates to inhibition of legumain or legumain related endopeptidase by cystatin. The inhibitory effect on legumain is by virtue of activity of a "second site" on cystatin, not previously known. The "first site" on cystatin is known to bind papain and papain-related peptidase. Papain does not bind the "second site" on cystatin.

Summary

We have investigated the inhibition of the recently identified family C13 cysteine peptidase, pig legumain, by human cystatin C. The cystatin was seen to inhibit enzyme activity by stochiometric 1:1 binding in competition with substrate. The K_i value for the interaction was 0.20 nM, i.e. cystatin C had an affinity for legumain similar to that for the papain-like family C1 cysteine peptidase, cathepsin B. However, cystatin C variants with alterations in the N-terminal region and the "second hairpin loop" that rendered the cystatin inactive against cathepsin B, still inhibited legumain with K_i values 0.2-0.3 nM. Complexes between cystatin C and papain inhibited legumain activity against Bz-Asn-$NHPhNO_2$ as efficiently as did cystatin C alone. Conversely, cystatin C inhibited papain activity against Bz-Arg-$NHPhNO_2$ whether or not the cystatin had been incubated with legumain, strongly indicating that the cystatin inhibited the two enzymes with non-overlapping sites. A ternary complex between legumain, cystatin C, and papain was demonstrated by gel filtration supported by immunoblotting. Screening of a panel of cystatin superfamily members showed that type 1 inhibitors (cystatins A and B) and L-kininogen (type 3) did not inhibit pig legumain. Of human type 2 cystatins, cystatin D was non-inhibitory, whereas cystatin E/M and cystatin F displayed strong (K_i 0.0016 nM) and relatively weak (K_i 10 nM) affinity for legumain, respectively. Sequence alignments and molecular modelling lead to the suggestion that a loop located on the opposite side to the papain-binding surface, between the α-helix and the first strand of the main β-pleated sheet of the cystatin structure, could be involved in legumain binding. This was corroborated by analysis of a cystatin C variant with substitution of the Asn39 residue in this loop (N39K-

cystatin C); this variant showed almost unaltered affinity for cathepsin B (K_i 1.5 nM) but >>5,000-fold lower affinity for legumain (K_i >>1,000 nM) than wildtype cystatin C.

The activities of cysteine peptidases of the papain family (C1) such as cathepsins B, H, L, S, and K in and around mammalian cells are regulated by reversible, tight-binding protein inhibitors of the cystatin superfamily (1). The cystatins constitute a superfamily of evolutionarily-related proteins that are all composed of at least one 100-120 residue domain with conserved sequence motifs (2). The single-domain human members of this superfamily are of two major types. The type 1 cystatins (or stefins) A and B contain approximately 100 amino acid residues, lack disulfide bridges, and are synthesized without signal peptides. Cystatins of type 2 are secreted proteins of approx. 120 amino acid residues (M_r 13,000-14,000) and contain at least two characteristic intrachain disulfide bonds. The type 2 cystatins include the human cystatins C, D, S, SN, and SA, which are all products of genes located in the cystatin multigene locus on chromosome 20 (3). Two recently identified type 2 cystatins, cystatin E/M and cystatin F (also called leukocystatin), are also secreted low-M_r proteins but are more atypical in that they are glycoproteins and show only 30-35% sequence identity in alignments with the classical type 2 cystatins. They are, however, still functional inhibitors of family C1 cysteine peptidases (4-7). It has been shown that the cystatin inhibiton of cysteine peptidases of the papain family is due to a tripartite wedge-shaped structure with very good complementarity to the active-site clefts of such enzymes (8). The three parts of the cystatin polypeptide chain included in the enzyme-binding domain are the N-terminal segment, a central loop-forming segment with the motif Gln-Xaa-Val-Xaa-Gly, SEQ ID No. 1, and a second C-terminal hairpin loop typically containing a Pro-Trp, SEQ ID No. 2, pair (8-10).

Legumain (EC 3.4.22.34) is a cysteine endopeptidase that was until recently known only from plants (11, 12) and *Schistosoma* (13). In plants there is abundant evidence that legumain performs a protein-processing function, causing limited proteolysis of precursor proteins and protein splicing (12, 14). Following the discovery of the enzyme in mammalian cells, it was cloned and sequenced from human (15) and mouse (16). The amino acid sequences of legumains show that they belong to a distinct family of cysteine endopeptidases (C13). Mammalian legumain is predominantly lysosomal in distribution (16), but its strict specificity for the hydrolysis of bonds on the carboxyl side of asparagine

is very different from that of any cathepsin and adapts it particularly for limited proteolysis (17). Human legumain may have an important physiological function as a key enzyme in antigen presentation (18).

It was recently reported that pig legumain is inhibited by human cystatin C and chicken cystatin with K_i values below 5 nM (15). This finding was unexpected, since the cystatins are already known as potent inhibitors of the papain-like cysteine peptidases in the unrelated family C1. The legumain family members are believed to have a protein fold quite unlike that of papain, and to be much more closely related to the caspases and gingipain (19). Although the active site cysteine residue could seem to be a common factor, it is known not to be required for the interaction of papain with cystatins (20). The present investigation was undertaken to elucidate the mechanism of inhibition of mammalian legumain by cystatins.

FIGURES

Figure 1 shows titration of legumain and papain activity by cystatin C.

Figure 2 shows analysis of the complex between cystatin C and legumain.

Figure 3 shows size exclusion chromatography of mixtures of cystatin C, legumain and papain.

Figure 4* shows sequence alignment of human members of the cystatin superfamily.

Figure 5* shows ribbon representation of the cystatin structure.

Figure 6* shows structural alignment of cystatins.

* full colour copy of these figures is available in M. Alvarez-Fernandez *et al*, 1999, "The Journal of Biological Chemistry", Vol. 274, No. 27, pp 19195-19203.

EXPERIMENTAL PROCEDURES

Proteins

Recombinant human cystatin D (the natural Arg26 variant) and wildtype human cystatin C were produced in *E. coli* and purified as earlier described (21, 22). Recombinant wildtype human cystatin E/M and F were isolated after baculovirus expression (4, 5). Human cystatin A and bovine cystatin B were obtained from Dr. I. Bjork (Uppsala, Sweden) as isolated *E. coli* produced proteins (23). Human L-kininogen was isolated from blood plasma (24). Papain (EC 3.4.22.2) was purchased from Sigma, St. Louis, MO (Cat. no. P4762) and used directly or after affinity purification on Sepharose coupled Gly-Gly-Tyr-Arg, which resulted in 70-75% activatable enzyme (25, 26). Papain was carboxymethylated using an earlier described procedure (27). Mammalian legumain (EC 3.4.22.34) was purified from pig kidney and characterized as recently reported (15). Human cathepsin B (EC 3.4.22.1) was obtained from Calbiochem (La Jolla, CA).

Production of cystatin C variants

Cystatin C devoid of the N-terminal 10 residues ((des 1-10)-cystatin C) was obtained by incubation of recombinant wildtype human cystatin C with neutrophil elastase and isolated by SEC (28). Cystatin C variants with Gly replacements for one, three or four of the residues involved in papain interactions, W106G-, (R8G,L9G,V10G)-, and (R8G,L9G,V10G,W106G)-cystatin C, were isolated after mutagenesis in an *E. coli* expression system (26). Dimeric cystatin C was obtained from wildtype recombinant human cystatin C by incubation for 30 min at 70 °C, and purified from trace amounts of monomeric cystatin C by SEC (29).

A cystatin C variant with Lys substitution for residue Asn39 was obtained by oligonucleotide-directed mutagenesis of the cystatin C cDNA gene in pHD313 (22, 30), using a PCR protocol. Taking advantage of a unique *Pst* I recognition site located 24 bp downstream from the Asn39 codon, a downstream primer was designed to introduce a C->A substitution in the Asn39 codon to result in a AAA codon for Lys (5'-CCACCTGCAGCGCGCGGCTGTGGTACATGTCTTTGCT-3'; SEQ ID No. 3 – strand, with *Pst* I site and mutation underlined). This oligonucleotide was used together with the

upstream vector primer MA206 (26), 5'-GTTTCGCCTGTCTGTTTTGC-3', SEQ ID No. 4 (both at 0.4 μM final concentration) to amplify a 360 bp fragment by PCR using 0.1 ng pHD313 DNA as template. DNA polymerase and PCR buffer were from the AmpliTaq kit (Perkin Elmer–Cetus, Norwalk, CT) and the PCR was accomplished by 30 incubation cycles of (95°C, 1 min – 57 °C, 1 min – 72 °C, 1 min) in a Perkin Elmer–Cetus 2400 thermocycler. The PCR product was purified (PCR Purification Kit; Genomed, Oeynhausen, Germany), digested with *Cla* I and *Pst* I (Life Technologies, Paisley, U.K.) and ligated into *Cla I/Pst* I cut and dephosphorylated pHD313, to generate plasmid pCmut39K. The plasmid was introduced into *E. coli* MC1061 as described in detail elsewhere (26). That the plasmid was correctly mutated was verified by complete nucleotide sequencing of the cystatin C insert, as described (26). The conditions for culturing and induction of expression in bacteria containing pCmut39K were as previously described for wildtype cystatin C production using pHD313 (22). Periplasmic extracts containing the recombinant cystatin C variant were obtained by cold osmotic shock (31) and directly applied to a Superdex 75 (Amersham Pharmacia Biotech, Uppsala, Sweden) SEC column (1.6 x 100 cm; in 50 mM ammonium bicarbonate buffer, pH 7.8, containing 100 mM NaCl). Fractions containing N39K-cystatin C were identified by agarose gel electrophoresis (32) and pooled. The purified cystatin variant was a homogenous protein preparation (>95% pure as estimated by SDS-polyacrylamide and agarose gel electrophoreses) with size and charge according to SDS-PAGE and agarose gel electrophoresis as expected.

Analysis of Cystatin C Complexes with Pig Legumain and Papain

Concentrations of human cystatin C and papain in solutions were determined by absorbance measurements, using ε_{280} values of 11,100 (33) and 58,500 (34) M^{-1} cm^{-1}, respectively. An ε_{280} value for pig kidney legumain of 47,100 M^{-1} cm^{-1} was calculated after quantitative amino acid analysis of a highly purified enzyme sample, by standard methods. The total cystatin and enzyme concentrations determined in this way are used in the text below, if not otherwise stated.

Cystatin C complexes with legumain were formed by incubating cystatin C with active pig legumain (from a 30 μM solution in 50 mM sodium citrate buffer, pH 5.5, containing

0.4 M NaCl, 1 mM EDTA, 0.01% (w/v) CHAPS, and 10 mM cysteine) in SEC buffer (below) for 30 min at room temperature. Mixtures analyzed by SEC (below) contained 186 μM cystatin C and 7.1 μM legumain. Ternary complexes between cystatin C and the two peptidases were typically formed by first incubating cystatin C with Cm-papain in SEC buffer during 30 min, and then adding active pig legumain and further incubating the mixture, at room temperature, for 30-60 min. Mixtures analyzed by SEC contained cystatin C, Cm-papain, and legumain at final concentrations of 81, 16, and 10 μM, or 9, 24, and 14 μM, respectively.

Separation and size estimation of the different enzyme-inhibitor complexes was performed by SEC on a Superdex 75 HR 10/30 column (Amersham Pharmacia Biotech) equilibrated in 50 mM sodium citrate buffer, pH 5.6, containing 150 mM NaCl. The column was operated at a flow rate of 0.5 mL/min using an HPLC system (Waters) equipped with multiple wavelength detector and an integration system (Waters 990). Ovalbumin (M_r 43,000), bovine serum albumin (M_r 67,000), chymotrypsinogen (M_r 23,400) and carbonic anhydrase (M_r 30,000) were used for construction of a calibration curve.

The fractions corresponding to each SEC peak were pooled and concentrated by precipitation, by addition of 9 volumes of 20% (w/v) trichloroacetic acid. Precipitated proteins in pellets obtained after centrifugation were resuspended in a minimal volume of SDS sample buffer, and analyzed by SDS-PAGE in 16.5% gels using the buffer system described by Schagger and von Jagow (35).

Immunoblotting

To verify the identity of protein bands after SDS-PAGE separation (above), transfer to PVDF membranes (Immobilon-P®; Millipore, Bedford, MA) was performed using electrophoresis (Trans-Blott®SD; Bio-Rad, Hercules, CA). Immunodetection of cystatin C was done exactly as described before (36). The same procedure was followed for Cm-papain detection with polyclonal rabbit anti-papain antibodies (produced by standard immunization procedures, using Cm-papain as antigen). Legumain was detected by use of horseradish peroxidase conjugated ConA (Sigma; Cat. no. L 6397) at a final concentration of 5 μg/mL. Antibody- or ConA-detected protein bands were visualized using

chemiluminescence (ECL Plus®; Amersham). Band intensities were assessed by densitometric scanning, using a Bio-Rad Imaging Densitometer GS-670 (Bio-Rad) and the Molecular Analyst software (Bio-Rad).

Enzyme Inhibition Assays

The methods used for active site titration of papain (with Bz-DL-Arg-NHPhNO_2 as substrate; Bachem Feinchemikalien, Bubendorf, Switzerland) and for titration of the molar papain-inhibitory concentration in cystatin preparations have been reviewed (1). Active inhibitor concentrations determined in this way were used for calculation of K_i values, as this is the method traditionally used. However, freshly isolated cystatin C preparations typically display apparent activities of 50-70% if the results from such papain titration assays are compared to total protein concentration determined by A_{280} measurement (22, 24). The apparently lower inhibitor concentration is likely due to to some of the papain molecules being catalytically inactive (possibly due to oxidation of the catalytic Cys residue) but still capable of binding cystatin. The stochiometry of papain–cystatin C binding is indeed 1:1 viewed by molar total protein concentrations, according to fluorescence titration (33). Therefore, total protein concentrations are used in the text when describing SEC experiments with papain–cystatin C mixtures.

The fluorogenic substrate used for determination of equilibrium constants for dissociation (K_i) of complexes between cystatins and family C1 cysteine peptidases (1) was Z-Phe-Arg-NHMec, SEQ ID No. 5 (10 μM; from Bachem Feinchemikalien) and the assay buffer was 100 mM Na-phosphate buffer (adjusted to pH 6.5 for papain, pH 6.0 for cathepsin B), containing 1 mM dithiothreitol and 1 mM EDTA. Steady state velocities were measured before and after addition of the cystatin variant under study in assays at 37 °C, and K_i values were calculated according to Henderson (37). Corrections for substrate competition were made using K_m values determined for the substrate batch used, under the assay conditions employed (60 and 55 μM for papain and cathepsin B, respectively).

Essentially the same procedures were used for legumain inhibition assays. Pig legumain was titrated with a cystatin C solution of known total protein concentration in a microtiter plate format using Bz-Asn-NHPhNO_2 as substrate (Bachem Feinchemikalien),

in sodium citrate/phosphate buffer (15), pH 5.8 (39.5 mM citric acid/121 mM Na_2HPO_4), containing 1 mM dithiothreitol, 1 mM EDTA, and 0.1% (w/v) CHAPS. The same buffer was used for fluorogenic continuous rate assays at 37 °C with 10 μM Z-Ala-Ala-Asn-NHMec, SEQ ID No. 6, prepared as described by Kembhavi et al. (11), as substrate. The legumain concentration used for K_i determinations in such assays was 0.1-0.5 nM. Results from such assays with substrate concentration in the range 5-50 μM were used to assess whether the cystatin interaction was competetive with substrate binding, by standard methods. The K_m value for legumain hydrolysis of this substrate under the assay conditions, used for corrections of apparent K_i values, was 30 μM.

Computer Modelling

Sequence alignments were carried out using programs in the GCG package (38), taking into account the structural alignment (39) of the known structures of chicken cystatin (Protein Data Bank (PDB) # 1CEW and PDB # 1A67 (8, 40)), cystatin A (PDB # 1DVC and PDB # 1DVD (41)), and cystatin B (PDB # 1STF (10)), all obtained from the Brookhaven Protein Data Bank (42, 43). Graphic illustrations were produced using the program Swiss-PdbViewer (44, 45) and then rendered with QuickDraw3D (Apple Computer, Inc.).

RESULTS

Cystatin C is a tight-binding legumain inhibitor – To clarify whether cystatins are efficient legumain inhibitors or not, the interaction between pig legumain and human cystatin C was initially investigated. Cystatin C was able to completely inhibit legumain activity against Bz-Asn-$NHPhNO_2$ in a time- and dose-dependent manner. Titration curves drawn from experiments with varying cystatin concentrations in such assays were linear (see Fig. 1 for example). As the concentrations of enzyme and inhibitor in the assay were in the order of 1 μM, this indicated a relatively tight complex between enzyme and inhibitor with a K_i value below 10 nM. Using Z-Ala-Ala-Asn-NHMec, SEQ ID No. 7, in different concentrations as substrate in a continuous-rate legumain assay at lower enzyme concentration (0.1-0.5 nM), it was observed that the cystatin C interaction with the enzyme was reversible and competing with substrate binding (results not shown). Corrected for substrate competition, the K_i value for the cystatin C – legumain complex was 0.20 nM.

Fig. 1. Titration of legumain and papain activity by cystatin C. A. Chromogenic assays for the activity of papain (7.0 μM active concentration determined by E-64 titration), with Bz-Arg-$NHPhNO_2$ as substrate, titrated with ■, cystatin C (7.5 μM total protein concentration); and □, the same cystatin C solution but preincubated with legumain (4.6 μM cystatin-binding sites, determined by titration as in Fig. 1B). **B.** Activity of legumain, assayed with Bz-Asn-$NHPhNO_2$ as substrate, titrated with **0**, cystatin C (7.5 μM total protein concentration), and, ● the same cystatin C solution premixed with papain (2.3 μM active concentration). Following incubation of enzyme(s) and inhibitor at concentrations given above for 30 min, the mixtures were diluted 10-fold in the substrate assays. Triangles show the activity in control tubes devoid of cystatin C, but containing dilutions of the same amounts of legumain and papain as in the cystatin C mixtures, in A and B, respectively. The data points shown are mean values from duplicate measurements.

Studies by SEC gave evidence for a 1:1 interaction between inhibitor and enzyme, with a faster-eluting peak (9.0 mL) appearing upon mixing of the proteins, corresponding to a M_r of 46,300 (Fig. 2A). This size agrees very well with a theoretical M_r value of 46-47,000 calculated as the sum of the M_r for cystatin C (13,343, from the amino acid sequence of recombinant cystatin C (30)) and the M_r for pig legumain according to SDS-PAGE of the glycosylated native enzyme (33,600 and 33,100 estimated from the SDS-

polyacrylamide gels in Figs. 3C and 2B, respectively). The retention volumes of 9.65 and 12.4 mL for the individual components on the calibrated SEC column (Fig. 2A, arrows) equaled M_r values of 35,300 and 13,400 for legumain and cystatin C, respectively. Analysis of the peak at 9.0 mL by SDS-PAGE agreed with the expected protein staining for legumain and cystatin C if they were present in equimolar amounts in a complex (Fig. 2B). In addition, the dissociated complex in the SDS gel demonstrated a cystatin C band size identical to that of the native inhibitor with no signs of degradation products, indicating that the inhibitor is not cleaved as a result of the enzyme interaction.

Fig. 2. Analysis of the complex between cystatin C and legumain. A. Isolated cystatin C and legumain were incubated as described under Experimental Procedures and SEC of the mixture on a Superdex 75 10/30 column at a flow rate of 0.5 mL/min was performed in a HPLC system. The arrows above the chromatogram indicate the elution volumes for individually analyzed legumain and cystatin C. **B.** Fractions corresponding to the peaks with elution volumes of 9.0 and 12.4 mL were concentrated by precipitation with trichloroacetic acid and analysed by SDS-PAGE in a 16.5% gel after reduction. The gel was stained with Serva-Blue. Lane *M*, protein markers, with M_r values indicated to the *left*. Lane *1*, recombinant cystatin C; Lane *2*, pig kidney legumain; Lane *3*, HPLC fraction corresponding to the elution peak at 9.0 mL; Lane *4*, fraction corresponding to elution peak at 12.4 mL.

The papain-inhibitory site on cystatin C is not responsible for legumain inhibition – To elucidate which parts of the cystatin structure are involved in binding and inhibition of mammalian legumain, the interactions of cystatin C variants with alterations in the N-terminal region and the second hairpin loop were studied (Table I). The (des1-10)-cystatin C variant, devoid of the N-terminal decapeptide as a result of neutrophil elastase cleavage and with seriously compromised affinities for cathepsins B, H, and L (28), showed the same affinity for legumain as native cystatin C. Three cystatin C variants with Gly replacements for up to four critical residues with side-chains participating in the high-affinity binding between the inhibitor and papain-like cysteine peptidases (26) also displayed virtually unaltered affinity for legumain. Additionally, dimerized cystatin C, which has been shown to lose papain-inhibitory activity completely and to be due to

intermolecular binding via the papain-inhibitory reactive site regions of two monomer units (46, 47), was essentially as efficient as the monomeric cystatin in the inhibition of legumain. Thus, the cystatin C surface responsible for the inhibition of papain-like enzymes seemed not to be involved in legumain binding.

The strict substrate specificity of legumain, with a requirement for an Asn residue in the P_1 position, allowed studies of the papain – cystatin C interaction in the presence of legumain by use of Bz-Arg-NHPhNO_2 as substrate (Fig. 1A). It was observed that the dose-dependent inhibition of papain by dilutions of a cystatin C solution was virtually unaffected when a portion of the same cystatin solution had been preincubated with an approx. equimolar amount of legumain (under conditions favoring stochiometric interactions between enzyme and inhibitor). Essentially identical cystatin C–enzyme mixtures could be analyzed for the presence of legumain-inhibitory sites, as papain showed very slow hydrolysis of the legumain substrate, Bz-Asn-NHPhNO_2 (Fig. 1B). The dose-dependent inhibition of legumain by dilutions of the cystatin C solution was largely unaffected when a portion of the same cystatin solution was preincubated with papain (at the highest concentration possible given the concentrations of the stock solutions used). In the fluorogenic legumain assay with Z-Ala-Ala-Asn-NHMec, SEQ ID No. 7, as substrate, cystatin C preincubated with papain in 1:1 and 1:10 molar ratio displayed K_i values of 0.26 and 0.33 nM, respectively, i.e. very similar to cystatin C alone (0.20 nM).

Demonstration of a ternary complex – Taken together, the results described above left little doubt that cystatin C inhibits legumain by a site that is distinct from that inhibiting papain and related peptidases. Consistent with this, the kinetic experiments indicated that the cystatin could simultaneously bind both legumain and papain, despite the small size of the cystatin molecule (M_r 13,343) and the two- and three-fold larger papain and legumain molecules, respectively. Attempts were therefore made to detect a ternary complex between the cystatin and the two peptidases. For experiments with excess papain, Cm-papain was used, to eliminate the risk of digestion of legumain by the papain. It was found that when a molar excess of cystatin C was incubated with Cm-papain and then with legumain, at final molar ratios of cystatin C:Cm-papain:legumain of 24:5:3, SEC analysis showed three peaks (Fig. 3A). Under these conditions, with a five-fold excess of cystatin C

over Cm-papain, most Cm-papain should become bound to the inhibitor and the major fraction of cystatin C would be free when legumain was added to the mixture. According to SDS-PAGE of the peak fractions (Fig. 3C), peaks III and II corresponded to free cystatin C and a Cm-papain–cystatin C complex, respectively. Peak I contained legumain and cystatin C in a bimolecular complex (*cf.* Fig. 2A). Next, cystatin C was incubated with Cm-papain and legumain in molar ratios of 2:5:3, to ascertain that most of the cystatin present in the mixture was bound to Cm-papain at the time of legumain addition. After incubation with legumain, SEC analysis of the mixture (Fig. 3B) demonstrated a new peak eluting at 8.7 mL (peak IV), earlier than that corresponding to the legumain-cystatin C di-complex, which corresponded to a M_r of approx. 53,000. The expected elution volume on the calibrated SEC column for a three-member complex with a theoretical M_r of 70,000 is 8 mL. SEC of Cm-papain alone or Cm-papain complexed with cystatin C resulted in anomalously low apparent M_r values, a phenomenon that has been seen repeatedly in the past (Barrett, A.J., unpublished observation). The anomalous behaviour of papain in SEC may well explain the small apparent size of a ternary complex in our SEC experiment. However, it was clear from SDS-PAGE analysis of the SEC peak IV at 8.7 mL (Fig. 3C) that it contained proteins with estimated M_r values of 33,600, 22,800, and 16,300, agreeing well with with pig legumain, Cm-papain, and cystatin C, respectively. The protein-staining of the three bands showed intensities consistent with a molar 1:1:1 ratio. Immunodetection of all three proteins after transfer to a PVDF membrane was performed on each fraction of the SEC shown in Fig. 3B. All three proteins were detected in the fractions included in peak IV, but were undetectable in the same fractions from control SEC experiments with cystatin C, legumain or Cm-papain alone (result not shown).

Fig. 3. Size exclusion chromatography of mixtures of cystatin C, legumain and papain. SEC was performed on a Superdex 75 column in a HPLC system as described under Experimental Procedures. The arrows above the chromatograms indicate the elution volumes for individually analyzed pig legumain, Cm-papain and recombinant human cystatin C. **A.** Cystatin C, Cm-papain, and legumain were incubated at a molar excess of cystatin C (molar ratio 24:5:3). **B.** Cystatin C, Cm-papain, and legumain were incubated at a limiting concentration of cystatin C (molar ratio approx. 2:5:3). **C.** The fraction corresponding to peak IV (8.7 mL) was concentrated by precipitation with trichloroacetic

acid and analyzed by SDS-PAGE after reduction, in a 16.5% gel stained with Serva-Blue. Lane *1*, recombinant human cystatin C; Lane *2*, Cm-papain; Lane *3*, pig kidney legumain; Lane *4*, peak IV. Lane *M*, protein markers, with M_r values indicated to the *left*.

Legumain-inhibitory activity of cystatin superfamily members – A panel of different cystatins was tested for inhibitory activity against pig kidney legumain. The results (Table 2) showed that the two type 1 inhibitors analyzed, human cystatin A and bovine cystatin B, had no significant inhibitory activity for pig legumain. Similarly, the type 3 inhibitor human L-kininogen, containing two cystatin repeats with inhibitory activity against papain-like cysteine peptidases, did not show any legumain-inhibitory activity. Of human type 2 cystatins, cystatins D was also non-inhibitory. Cystatin E/M demonstrated tight-binding affinity for pig legumain, with a K_i value almost 100-fold lower than that of cystatin C, whereas cystatin F showed a significant but lower affinity for the enzyme, 50-fold lower than that of cystatin C.

Possible location of the legumain inhibitory site on cystatins – An alignment of the amino acid sequences for the different cystatins (Fig. 4) was inspected for similarities between the proteins showing activity against legumain. Of those segments known to be present on the surfaces of the cystatin molecules (8, 10, 40, 41, 47), the loop starting at Asn39, following the conserved α-helix, seemed like a good candidate for inclusion in a legumain-binding surface. This was because: (1) it is located on the opposite side to the papain-inhibitory surface (Fig. 5), which would allow the simultaneous binding of legumain and papain observed for cystatin C; and (2) it contains residues that are conserved among the cystatins with legumain-inhibitory activity, including an Asn residue, which could be directly involved in interactions with the substrate specificity pocket of legumain, given the strict specificity of legumain for asparaginyl bonds (17). To try to verify the possibility that the Asn39 residue is intimately involved in the legumain-inhibitory site, this residue was mutated in cystatin C. The substitution selected was Asn39Lys, guided by the notion that the non-inhibitory cystatin D, as well as cystatin B, have a positively charged residue in the loop segment. The N39K cystatin C variant was produced by *E. coli* expression and purified to homogeneity. The variant showed a 5-fold

decreased affinity for cathepsin B as compared to wildtype cystatin C (Table I), but total loss of inhibitory activity for pig legumain (K_i, >>1,000 nM, equalling >>5,000-fold lower legumain affinity than wildtype cystatin C). This strongly indicated that the loop between the α-helix and the first strand of the main β-pleated sheet of the cystatin structure, and its Asn39 residue, is part of a novel second reactive site of some cystatins.

Fig. 4. Sequence alignment of human members of the cystatin superfamily. The alignment was done based on secondary structure elements from the NMR models of cystatin A and chicken cystatin (with structural elements indicated above the sequences of type 1 and type 2 cystatins, respectively) and the X-ray model of chicken cystatin (not represented in the figure) (8, 39-41). *Arrows* represent β-strands (type 1 and 2) and *cylinders* represent α-helices. The region possibly involved in a legumain-inhibitory site (the 'back-side loop') is shown. This loop-forming segment is magnified below; residues with similar chemical properties present in the three cystatins showing legumain-inhibitory activity (cf. Table II) but not in the others, and thus possibly constituting a consensus sequence for legumain binding, are *boxed*. Residue Asn39 for cystatin A and in cystatin C are also shown in Fig. 6. The numbering refers to the cystatin C sequence as deduced from its cDNA, starting from the first residue of the mature protein (30, 48). For the other cystatins, the naturally occurring forms with longest N-terminal segments are shown (4, 5, 21, 49-51). *Dots* indicate gaps introduced to optimize the alignment.

Fig. 5. Ribbon representation of the cystatin structure. The locations of amino acids substituted in the cystatin C variants studied are shown and labelled. The illustration is based on the NMR model for chicken cystatin (40). The region constituting the cystatin site responsible for inhibition of papain-like cysteine peptidases is marked. The main atoms of the wildtype Asn residue in position 39 are shown in a ball-and-stick representation.

DISCUSSION

The aim of the present investigation was to study the mechanism of inhibition of mammalian legumain by cystatins, to clarify how the inhibitor structure can result in tight-binding inhibition of enzymes belonging to two entirely different enzyme families, namely the papain family (C1) and the legumain family (C13). The different arrangements of catalytic residues and different active site motifs show that the two families are evolutionarily unrelated, and that their peptidases have different protein folds (19). Moreover, legumain is not inhibited by the general inhibitor of enzymes belonging to family C1, E-64 (15), which supports that the general topography of the active site clefts of legumains likely are entirely different from those of family C1 enzymes. Our present SEC, electrophoretic and enzyme kinetic results show that cystatin C can inhibit mammalian legumain, as cystatins inhibit family C1 enzymes (52), by high-affinity reversible binding (K_i 0.20 nM), in a bimolecular reaction that is competetive with substrate, and with no detectable cleavage of the cystatin in the legumain complex. Despite these similarities, our present results demonstrate that the mechanisms of inhibition of legumain and family C1 endopeptidases must be completely different.

From structural studies of several cystatins (8-10), it is well known that the N-terminal segment together with the 'first and second hairpin loops' in cystatins are responsible for the inhibition of the C1 enzymes (Fig. 5). Consequently, removal of the N-terminal segment or substitution of any of the conserved amino acids in the N-terminal segment or the hairpin loops by Gly/Ala residues abolishes or seriously affects the inhibition of papain-like enzymes (26, 53). Of four such variants analyzed in the present study, all displayed virtually unaltered binding of legumain. Dimeric cystatin C, which is completely inactive against papain-like enzymes and by NMR studies has been shown to be a result of intermolecular interactions between the papain-binding surfaces of two cystatin C molecules (47), still showed legumain inhibition. We believe that this, together with the enzyme kinetic results presented and the direct demonstration of a ternary complex by SEC, proves that the binding sites for papain and legumain on cystatin C likely are completely independent of each other.

Where then is the legumain reactive site? Our investigation of a set of other mammalian cystatin superfamily members indicated that the capacity to inhibit legumain is a property of only some cystatins (Table 2). Guided by this result and amino acid sequence comparisons, we propose that the side of cystatins directly opposite to the papain-binding surface is responsible for the legumain binding and inhibition. The loop segment connecting the main α-helix of the cystatin structure to the first long β-strand contains a conserved Asn residue (residue 39 in cystatin C) and seems quite conserved in sequence in those cystatins that show inhibitory activity: cystatins C, E/M, and F. An importance of the Asn39 residue was confirmed by construction of the N39K cystatin C variant, which was seen to lack legumain inhibitory activity. A correctly positioned Asn residue on the cystatin surface could possibly result in an initial substrate-like interaction between the inhibitor and legumain. Besides the requirement for an Asn residue in the P_1 position, legumains have no clear preferences for residues in other subsites (17, 54). There are therefore few obvious structural possibilities for specific legumain inhibition besides interaction with the S_1 pocket. Still, assuming that the 'back-side loop' containing Asn39 interacts with the enzyme in a mode resembling substrate binding, it appeared from the inhibition data obtained that a loop segment preferentially containing polar amino acids is compatible with legumain interaction. The consensus sequence found in the three inhibitory cystatins is Ser(Thr)-Asn39-Asp(Ser)-Met(Ile), SEQ ID No. 8, Strikingly, a Ser38-Asn39-Asp40, SEQ ID No. 9, sequence is completely conserved in mouse, rat, and bovine cystatin C, as well as in chicken cystatin, which also inhibits pig legumain (15). The positively charged Lys residue in this segment, present in the non-inhibitory cystatins B and D, may be unfavourable for inhibition.

The suggested binding loop must be able to adopt a conformation to allow legumain interaction, but at the same time not expose the Asn39-Xaa bond to cleavage. A different 'back-side loop' conformation may be one reason why the type 1 cystatins studied, with a loop sequence largely containing the proposed consensus sequence for legumain inhibition, although being two residues smaller, do not show inhibitory activity (Fig. 6). In the case of the inhibitory cystatins, the loop might be partially restrained in cystatin F as Cys37 likely is involved in a disulfide bridge (5), which can explain why cystatin F is a poorer inhibitor than cystatin C and E/M. The size and conformation of the loop could also be one reason

why cystatin D does not inhibit legumain, because of an amino acid insertion in this loop (Fig. 4). For the type 3 cystatin studied, human L-kininogen, the lack of legumain-inhibitory activity may be due to sterical reasons, as both legumain and kininogen are bulky molecules. Two of the three cystatin domains of kininogens are clearly able to inhibit papain-like peptidases (55), which demonstrates that the papain-binding surfaces of these domains are exposed and accessible to protein interactions. Whether the kininogen structure is sufficiently flexible to also allow exposure of the 'back-side loops' on the opposite sides of these domains is presently unclear, as a three-dimensional model for type 3 cystatins is unfortunately not yet available. For the individual kininogen domains, the sequence requirements for a legumain-binding 'back-side loop' suggested above seem to be fulfilled for domain 3, but not for domain 2 of human kininogen. Clearly, more studies are needed to clarify whether perhaps some variants of L- or H-kininogens, resulting from proteolytic cleavages to release the kinin portion or individual cystatin domains of the protein, display legumain-inhibitory activity.

Fig. 6. Structural alignment of cystatins. A three-dimensional alignment of cystatin A and chicken cystatin, zoomed in on the 'back-side loop' region close to the cystatin α-helix, is shown in two different orientations. The loop and part of the first long β-strand of cystatin A (41) is shown as (a) and the corresponding segment of chicken cystatin (40) as (b). **A**, Side-view with the orientation of the chains from the N-terminal ends of the α-helices in the *upper left* corner. The alignment demonstrates that not only Asn39 (as (c)) but the entire loop is more exposed at the molecule surface in chicken cystatin when compared to the corresponding residue in cystatin A (as (d)), which is placed in a considerably shorter loop. **B**, View along the α-helices, from their C-terminal ends. The residues corresponding to Asn39 in cystatin C are shifted by 5.2Å from each other, probably as a result of the kink in the third turn of the α-helix in cystatin A.

Although our initial studies indicate that the 'back-side loop' around Asn39 is important for the ability of some cystatins to efficiently inhibit legumain, other cystatin segments may also be involved in interactions with the enzyme, just as several segments are involved in the cystatin inhibition of papain. The very flexible loop between the 2nd and 3rd of the four main β-strands of the cystatin structure, from Thr74 to Asn82 (which is

not present in type 1 cystatins) may prove essential to stabilize the enzyme–inhibitor interaction, given its close proximity to the Asn39 loop (Fig. 5). Interestingly, this segment contains a five-residue insertion in the most efficient legumain inhibitor we identified, cystatin E/M (4, 6). That this loop contains the primary binding site for legumain seems quite unlikely, however, as the loop sequence is relatively conserved between human cystatins C and D (Fig. 4), of which only cystatin C shows legumain-inhibitory activity.

In conclusion, our present results strongly indicate that the loop between the α-helix and the first strand of the main β-pleated sheet of the cystatin structure and its Asn39 residue, is part of a novel second reactive site of some cystatins. Cystatins carrying this site are sufficiently potent to be physiological inhibitors of mammalian legumain. Since legumain-like activity has very recently been shown to be crucial for cellular presentation of certain antigens to the immune system, but no efficient inhibitors to this activity are presently known (18), continued studies to elucidate and explore the mechanism of legumain inhibition by the novel cystatin site may prove valuable.

Table I

Inhibition of legumain by cystatin C variants

Equilibrium constants for dissociation of pig legumain and human cathepsin B complexes with cystatin C variants were determined by steady-state kinetics in continuos rate assays at 37 °C with Z-Ala-Ala-Asn-NHMec, SEQ ID No. 6, and Z-Phe-Arg-NHMec, SEQ ID No. 5, respectively, as substrate. The standard deviation (S.D.) and number of measurements (n) used to calculate the mean K_i values given are indicated. The K_i values were corrected for substrate competition as described under Experimental Procedures.

Cystatin C Variant	Legumain			Cathepsin B		
	K_i (nM)	S.D.	n	K_i (nM)	S.D.	n
Wildtype	0.20	±0.014	(6)	0.25	-	(1)
(des1-10)-	0.26	±0.056	(4)	>>1000[a]		
(R8G,L9G,V10G,W106G)-	0.22[b]	±0.011	(3)	>>1000[c]		
Dimeric[d]	0.39	±0.027	(2)	>>1000[e]		
N39K-	>>1000[f]			1.5	±0.14	(3)

[a] $v_i/v_0 > 0.95$ at [I] = 670 nM

[b] Results for W106G- and (R8G,L9G,V10G)-cystatin C were, as that for the (R8G,L9G,V10G,W106G) variant, not significantly different from the wildtype cystatin K_i value.

[c] $v_i/v_0 > 0.95$ at [I] = 380 nM

[d] Calculated from active papain-inhibitory concentration of monomeric cystatin C, determined prior to dimerization, and assuming no protein loss during the following SEC purification.

[e] $v_i/v_0 > 0.95$ at [I] = 510 nM

[f] $v_i/v_0 > 0.95$ at [I] = 396 nM

Table II

Inhibition of legumain by cystatins of type 1, 2 and 3

Equilibrium constants for dissociation of pig legumain complexes with recombinant bovine cystatin B, recombinant human cystatins A, C, D, E/M, and F, and low-M_r kininogen purified from human blood plasma were determined by steady-state kinetics in continuos rate assays at 37 °C with Z-Ala-Ala-Asn-NHMec, SEQ ID No. 6, as substrate. The standard deviation (S.D.) and number of measurements (n) used to calculate the mean K_i values given are indicated. The K_i values were corrected for substrate competition as described under Experimental Procedures. v_0, rate of substrate hydrolysis in absence of inhibitor; v_i, rate of substrate hydrolysis in presence of inhibitor. The quality of the cystatin preparations used was checked by determination of K_i values for their interaction with papain or cathepsin B (published values are reviewed in Refs. 1, 5).

	Legumain			Cystatin activity control
	K_i (nM)	S.D.	n	
Cystatin A	>>1000[a]			Cathepsin B: K_i = 1.5±0.58 nM (n=2)
Cystatin B	>>1000[b]			Papain: K_i = 0.037±0.010 nM (n=4)
Cystatin C	0.20	±0.014	(6)	Cathepsin B: K_i = 0.25 nM (n=1)
Cystatin D	>>1000[c]			Papain: K_i = 1.9±1.4 nM (n=4)
Cystatin E/M	0.0016	±0.00052	(5)	Papain: K_i = 0.39±0.13 nM (n=12)
Cystatin F	10	±1.0	(4)	Papain: K_i = 1.1±0.26 nM (n=5)
L-kininogen	>>1000[d]			Papain: K_i = 0.030±0.012 (n=3)

[a] $v_i/v_0 > 0.95$ at [I] = 670 nM

[b] $v_i/v_0 > 0.95$ at [I] = 688 nM

[c] $v_i/v_0 > 0.95$ at [I] = 2120 nM

[d] $v_i/v_0 > 0.95$ at [I] = 780 nM

REFERENCES

1. Abrahamson, M. (1994) *Methods Enzymol.* **244**, 685-700

2. Rawlings, N. D., and Barrett, A. J. (1990) *J. Mol. Evol.* **30**, 60-71

3. Schnittger, S., Rao, V. V., Abrahamson, M., and Hansmann, I. (1993) *Genomics* **16**, 50-55

4. Ni, J., Abrahamson, M., Zhang, M., Alvarez-Fernandez, M., Grubb, A., Su, J., Yu, G.-L., Li, Y., Parmelee, D., Xing, L., Coleman, T. A., Gentz, S., Thotakura, R., Nguyen, N., Hesselberg, M., and Gentz, R. (1997) *J. Biol. Chem.* **272**, 10853-10858

5. Ni, J., Alvarez-Fernandez, M., Danielsson, L., Chillakuru, R. A., Zhang, J., Grubb, A., Su, J., Gentz, R., and Abrahamson, M. (1998) *J. Biol. Chem.* **273**, 24797-24804

6. Sotiropoulou, G., Anisowicz, A., and Sager, R. (1997) *J. Biol. Chem.* **272**, 903-910

7. Halfon, S., Ford, J., Foster, J., Dowling, L., Lucian, L., Sterling, M., Xu, Y., Weiss, M., Ikeda, M., Liggett, D., Helms, A., Caux, C., Lebecque, S., Hannum, C., Menon, S., McClanahan, T., Gorman, D., and Zurawski, G. (1998) *J. Biol. Chem.* **273**, 16400-16408

8. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988) *EMBO J.* **7**, 2593-2599

9. Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W., and Barrett, A. J. (1987) *J. Biol. Chem.* **262**, 9688-94

10. Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., and Turk, V. (1990) *EMBO J.* **9**, 1939-1947

11. Kembhavi, A. A., Buttle, D. J., Knight, C. G., and Barrett, A. J. (1993) *Arch. Biochem. Biophys.* **303**, 208-213

12. Hara-Nishimura, I. (1998) in *Handbook of Proteolytic Enzymes* (Barrett, A. J., Rawlings, N. D., and Woessner, J. F., eds) pp.746-749, Academic Press, London, U. K.

13. Dalton, J. P., and Brindley, P. J. (1998) in *Handbook of Proteolytic Enzymes* (Barrett, A. J., Rawlings, N. D., and Woessner, J. F., eds) pp.749-754, Academic Press, London, U. K.

14. Min, W., and Jones, H. (1994) *Structural Biology* **1**, 502-504

15. Chen, J.-M., Dando, P. M., Rawlings, N. D., Brown, M. A., Young, N. E., Stevens, R. A., Hewitt, E., Watts, C., and Barrett, A. J. (1997) *J. Biol. Chem.* **272**, 8090-8098

16. Chen, J. M., Dando, P. M., Stevens, R. A. E., Fortunato, M., and Barrett, A. J. (1998) *Biochem. J.* **335**, 111-117

17. Dando, P. M., Fortunato, M., Smith, L., Knight, C. G., McKendrick, J. E., and Barrett, A. J. (1999) *Biochem. J.*, in press

18. Manoury, B., Hewitt, E. C., Morrice, N., Dando, P. M., Barrett, A. J., and Watts, C. (1998) *Nature* **396**, 695-699

19. Chen, J.-M., Rawlings, N. D., Stevens, R. A. W., and Barrett, A. J. (1998) *FEBS Lett.* **441**, 361-365

20. Björk, I., and Ylinenjarvi, K. (1989) *Biochem. J.* **260**, 61-8

21. Freije, J. P., Balbín, M., Abrahamson, M., Velasco, G., Dalbøge, H., Grubb, A., and López-Otín, C. (1993) *J. Biol. Chem.* **268**, 15737-15744

22. Abrahamson, M., Dalboge, H., Olafsson, I., Carlsen, S., and Grubb, A. (1988) *FEBS Lett.* **236**, 14-18

23. Pol, E., Olsson, S. L., Estrada, S., Prasthofer, T. W., and Björk, I. (1995) *Biochem. J.* **311**, 275-282

24. Abrahamson, M., Barrett, A. J., Salvesen, G., and Grubb, A. (1986) *J. Biol. Chem.* **261**, 11282-11289

25. Blumberg, S., Schechter, I., and Berger, A. (1970) *Eur. J. Biochem.* **15**, 97-102

26. Hall, A., Håkansson, K., Mason, R. W., Grubb, A., and Abrahamson, M. (1995) *J. Biol. Chem.* **270**, 5115-5121

27. Anastasi, A., Brown, M. A., Kembhavi, A. A., Nicklin, M. J. H., Sayers, C. A., Sunter, D. C., and Barrett, A. J. (1983) *Biochem. J.* **211**, 129-138

28. Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A., and Ohlsson, K. (1991) *Biochem. J.* **273**, 621-626

29. Gerhartz, B., Ekiel, I., and Abrahamson, M. (1998) *Biochemistry* **37**, 17309-17317

30. Abrahamson, M., Grubb, A., Olafsson, I., and Lundwall, A. (1987) *FEBS Lett.* **216**, 229-33

31. Dalbøge, H., Jensen, E. B., Tottrup, H., Grubb, A., Abrahamson, M., Olafsson, I., and Carlsen, S. (1989) *Gene* **79**, 325-332

32. Jeppsson, J. O., Laurell, C. B., and Franzén, B. (1979) *Clin. Chem.* **25**, 629-638

33. Lindahl, P., Abrahamson, M., and Björk, I. (1992) *Biochem. J.* **281**, 49-55

34. Glazer, A. N., and Smith, E. L. (1971) in *The Enzymes* (Boyer, P., ed) pp.501-546, Academic Press, New York

35. Schägger, H., and von Jagow, G. (1987) *Anal. Biochem.* **166**, 368-379

36. Bjarnadottir, M., Wulff, B., Sameni, M., Sloane, B. F., Keppler, D., Grubb, A., and Abrahamson, M. (1998) *Mol. Pathol.* **51**, 317-326

37. Henderson, P. J. (1972) *Biochem. J.* **127**, 321-333

38. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395

39. Turk, V., and Bode, W. (1991) *FEBS Lett.* **285**, 213-9

40. Dieckmann, T., Mitschang, L., Hofmann, M., Kos, J., Turk, V., Auerswald, E. A., Jaenicke, R., and Oschkinat, H. (1993) *J. Mol. Biol.* **234**, 1048-1059

41. Martin, J. R., Craven, C. J., Jerala, R., Kroon-Zitko, L., Zerovnik, E., Turk, V., and Waltho, J. P. (1995) *J. Mol. Biol.* **246**, 331-343

42. Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., and Weng, J. (1987) in *Crystallographic Databases-Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., and Sievers, R., eds) pp.107-132, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester

43. Abola, E. E., Sussman, J. L., Prilusky, J., and Manning, N. O. (1997) *Methods Enzymol.* **277**, 556-571

44. Guex, N. (1996) *Experimentia* **52**, A26

45. Guex, N., and Peitsch, M. C. (1996) *Protein Data Bank Quaterly Newsletter* **77**, 7

46. Ekiel, I., and Abrahamson, M. (1996) *J. Biol. Chem.* **271**, 1314-21

47. Ekiel, I., Abrahamson, M., Fulton, D. B., Lindahl, P., Storer, A. C., Levadoux, W., Lafrance, M., Labelle, S., Pomerleau, Y., Groleau, D., LeSauteur, L., and Gehring, K. (1997) *J. Mol. Biol.* **271**, 266-277

48. Grubb, A., and Löfberg, H. (1982) *Proc. Natl. Acad. Sci. U.S.A* **79**, 3024-3027

49. Ritonja, A., Machleidt, W., and Barrett, A. J. (1985) *Biochem. Biophys. Res. Commun* **131**, 1187-1192

50. Machleidt, W., Borchart, U., Fritz, H., Brzin, J., Ritonja, A., and Turk, V. (1983) *Hoppe Seylers Z. Physiol. Chem.* **364**, 1481-6

51. Ohkubo, I., Kurachi, K., Takasawa, T., Shiokawa, H., and Sasaki, M. (1984) *Biochemistry* **23**, 5691-5697

52. Nicklin, M. J. H., and Barrett, A. J. (1984) *Biochem. J.* **223**, 245-253

53. Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. A., and Fritz, H. (1992) *Eur. J. Biochem.* **209**, 837-845

54. Abe, Y., Shirane, K., Yokosawa, H., Matsushita, H., Mitta, M., Kato, I., and Ishii, S. (1993) *J. Biol. Chem.* **268**, 3525-3529

55. Salvesen, G., Parkes, C., Abrahamson, M., Grubb, A., and Barrett, A. J. (1986) *Biochem. J.* **234**, 429-434

Claims

1. A process for inhibition of the enzymatic activity of legumain or a legumain-related endopeptidase by cystatin; which process does not involve interaction with a first papain-reactive site of cystatin; but which involves an interaction with a second papain-non-reactive site of cystatin, and wherein said process for inhibition can take place in the presence or absence of binding of papain or a related peptidase to the papain-reactive site of cystatin.
2. Use of an inhibitor comprising a papain-non-reactive site of cystatin to inhibit the enzymatic activity of legumain or a legumain-related endopeptidase whereby either;
   a) the papain reactive site of cystatin is not present; or
   b) the papain reactive site of cystatin is blocked by papain or a related peptidase.
3. Use of the non-papain-reactive but legumain-reactive site of cystatin in a method of modelling an inhibitor of the enzymatic activity of legumain or a legumain-related endopeptidase.
4. Use of the non-papain-reactive but legumain-reactive site of cystatin in a method to assay for potential cystein proteases, which potential proteases bind to said site.
5. Use according to claim 2 modified in that the non-papain reactive site of cystatin has a mutation at one or more residue within said site whereby said site can bind to and inhibit a class of cystein proteases structurally related to legumain; such as caspases, gingipain or clostripain.
6. A process or use according to any one of claims 1 to 5 wherein the papain-non-reactive site of cystatin comprises the loop between the α-helix and the first strand of the main β-pleated sheet of the cystatin structure.
7. A process or use according to any preceding claim wherein said papain-non-reactive site includes Asn39.
8. A process or use according to any preceding claim wherein said papain-non-reactive site is derived from a type 2 cystatin; such as human cystatin C or chicken cystatin.
9. An inhibitor of the enzymatic activity of legumain or a structurally related cystein protease, which inhbitor does not bind papain or papain-related peptidase.

10. A method of inhibiting a cystein protease in a numan or animal, which method includes the step of administering an inhibitor according to claim 9 in the form of a soluble protein or by gene therapy.

Figure 1

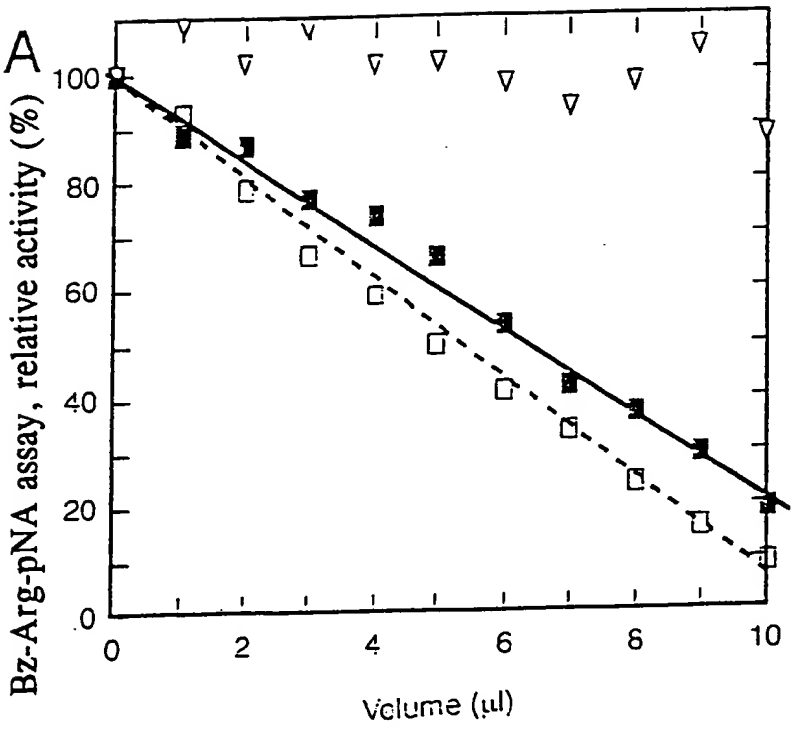

A
Bz-Arg-pNA assay, relative activity (%)
100
80
60
40
20
0
0
2
4
6
8
10
Volume (µl)

Figure 1 continued

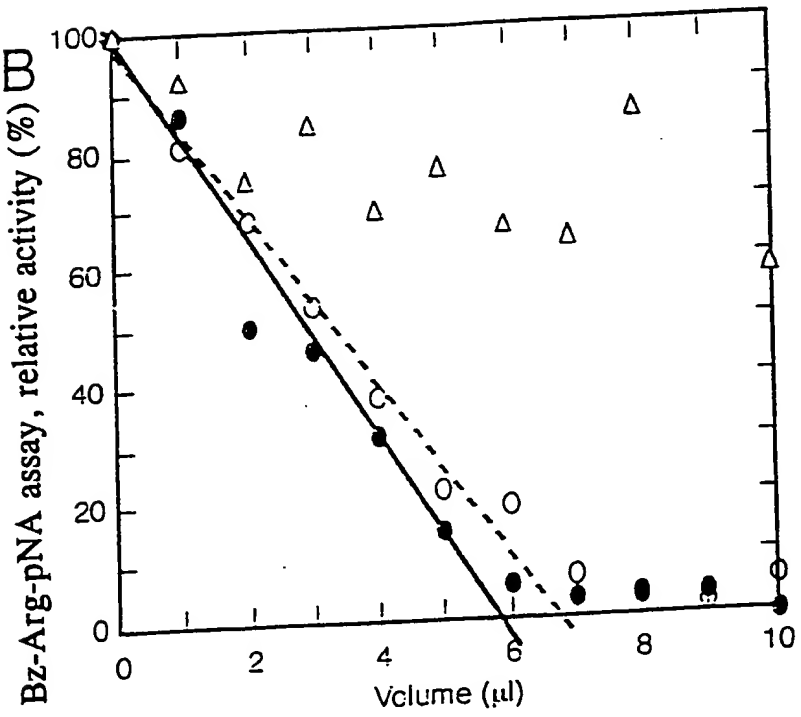


Fig. 2

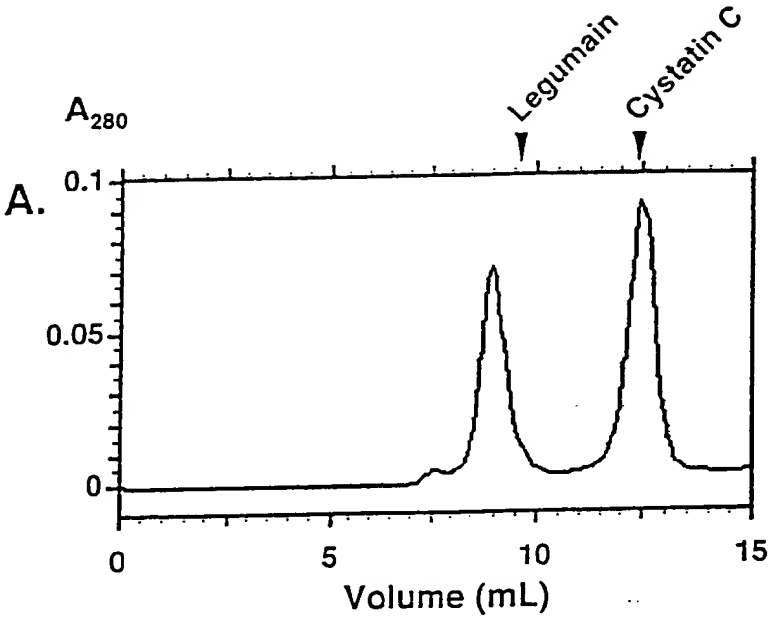


Figure 2 continued

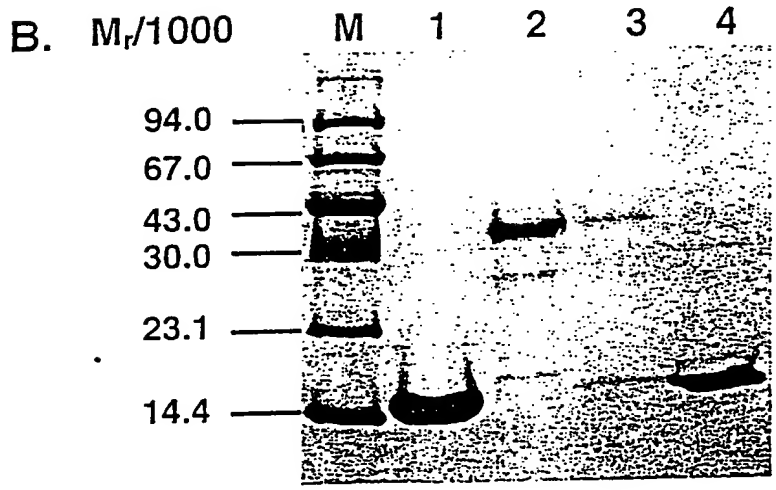


Figure 3

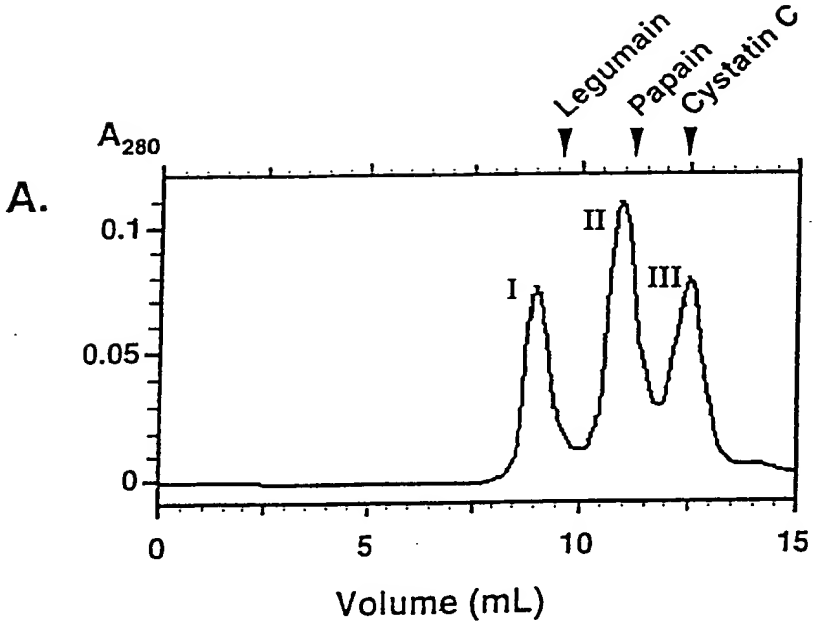

A.
Legumain
Papain
Cystatin C
A_{280}
0.1
0.05
0
I
II
III
0
5
10
15
Volume (mL)

Figure 3 continued

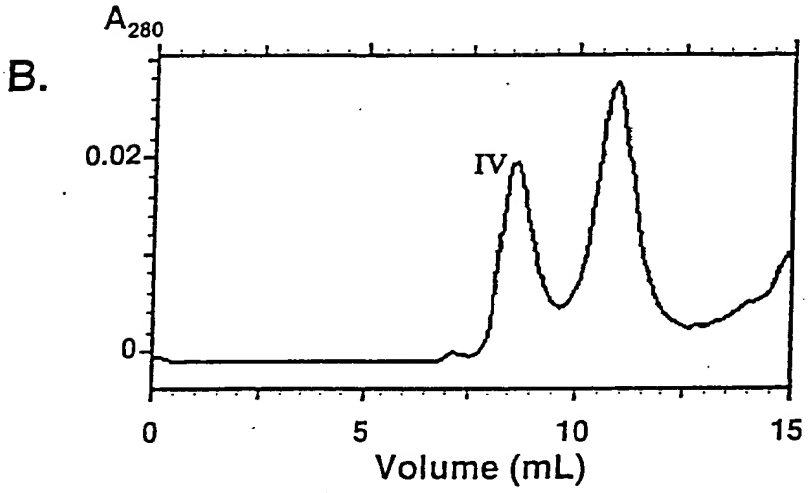
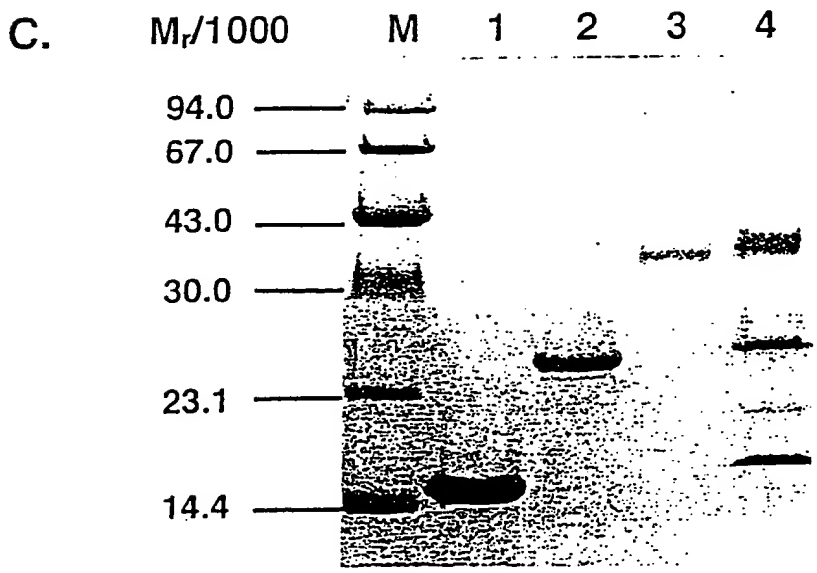


Figure 3 continued



1
10
20
Type 1
Cystatin A .....MIP GGLSEAKPAT
Cystatin B .....MMC GGTSATQPAT
Type 2
Cystatin C .....SSPGKPPRLV GGPMDASVEE
Cystatin D .....AQSRTLA GGIHATDLND
Cystatin E/M .....RPQERMV GELRDLSPDD
Cystatin F GPSPDTCSQDLNSRVKP GFPKTIKTND
Type 3
Kininogen D2 .....TAQYDCL GCVHPISTQS
Kininogen D3 .....PPTKICV GCPRDIPTNS

Figure 4

Figure 4 continued

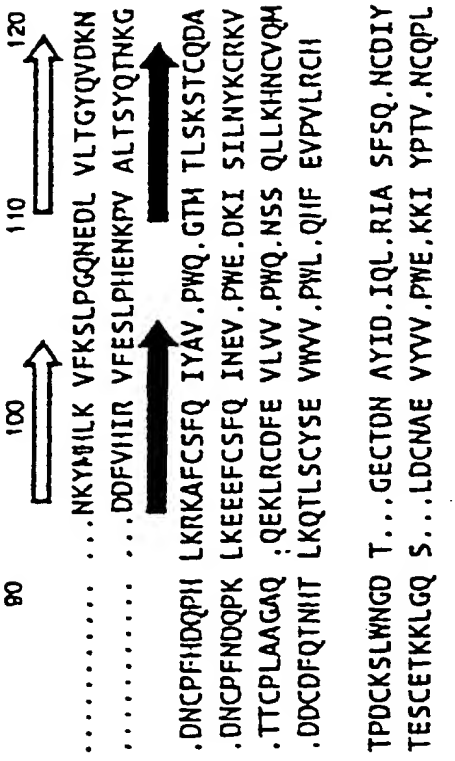


Figure 4 continued

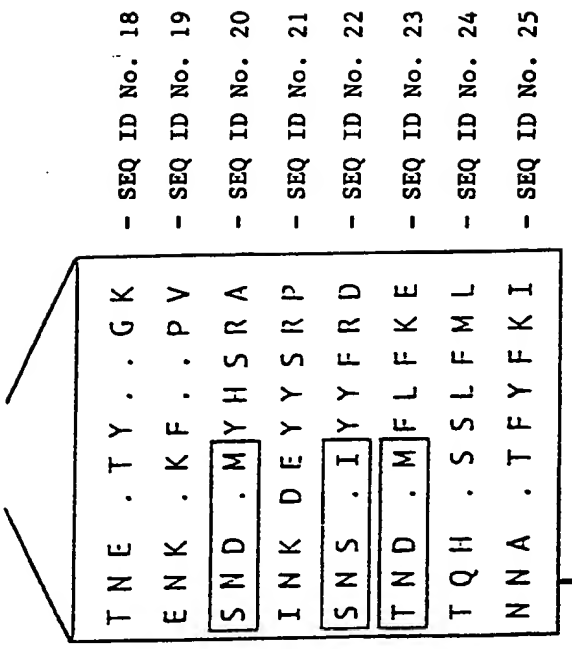


Figure 4 continued

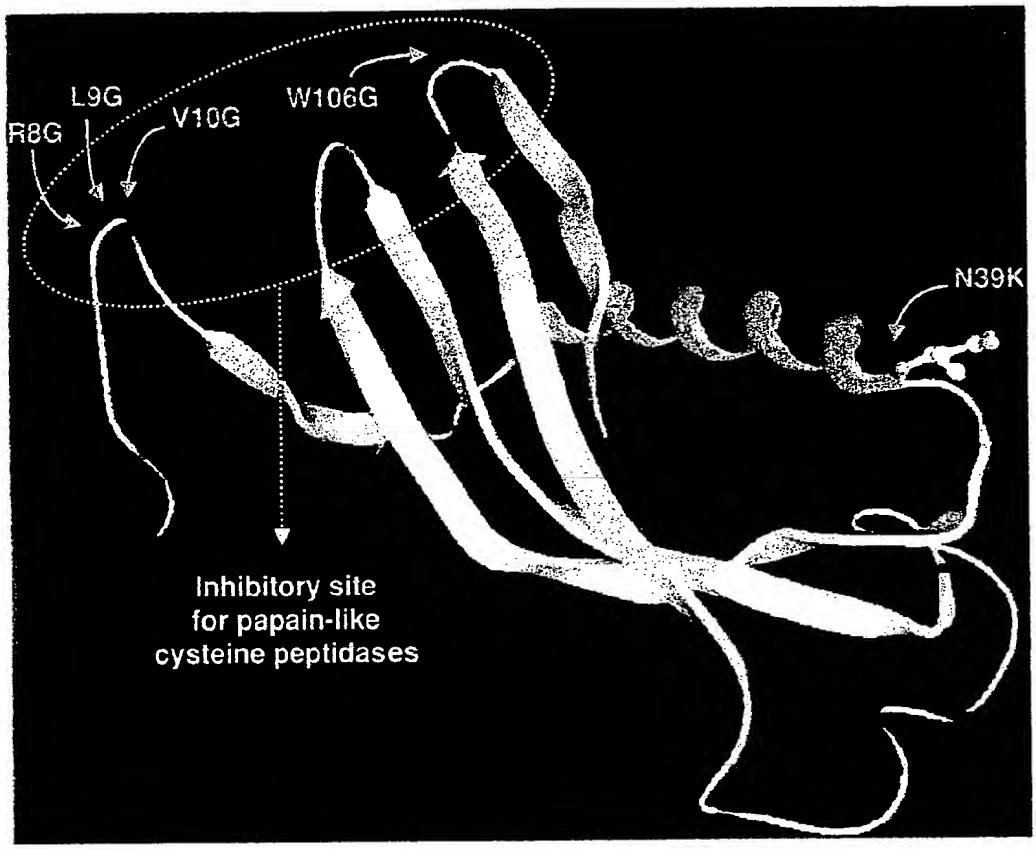

L9G
R8G
V10G
W106G
N39K
Inhibitory site
for papain-like
cysteine peptidases

Figure 5

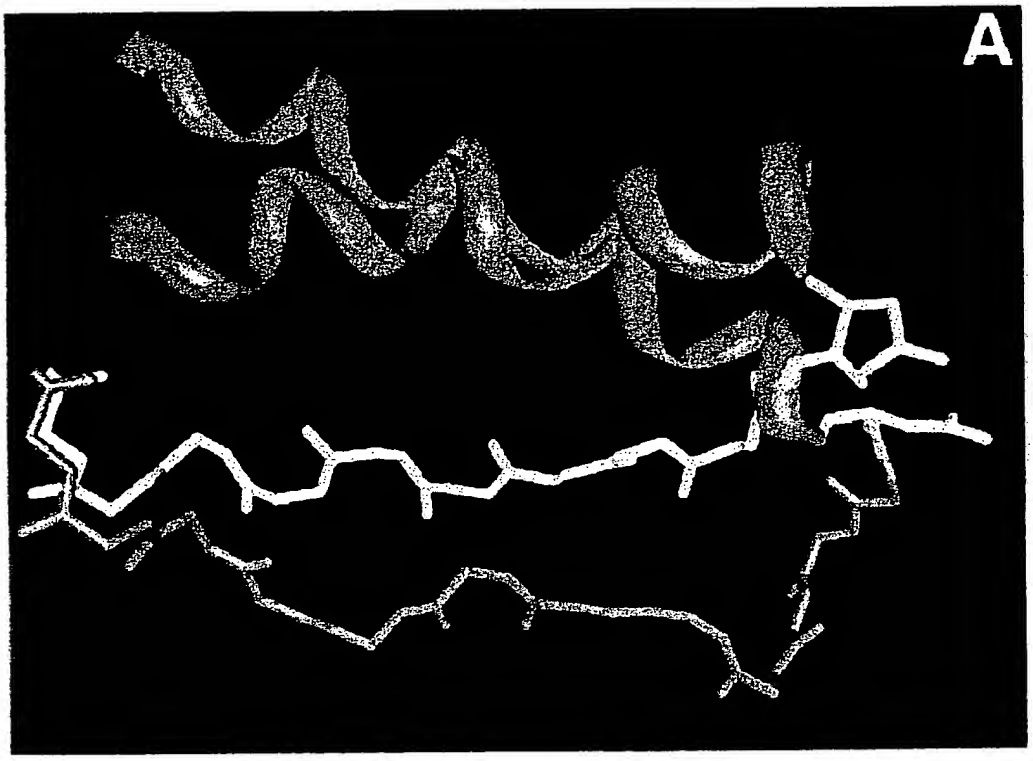


Figure 6

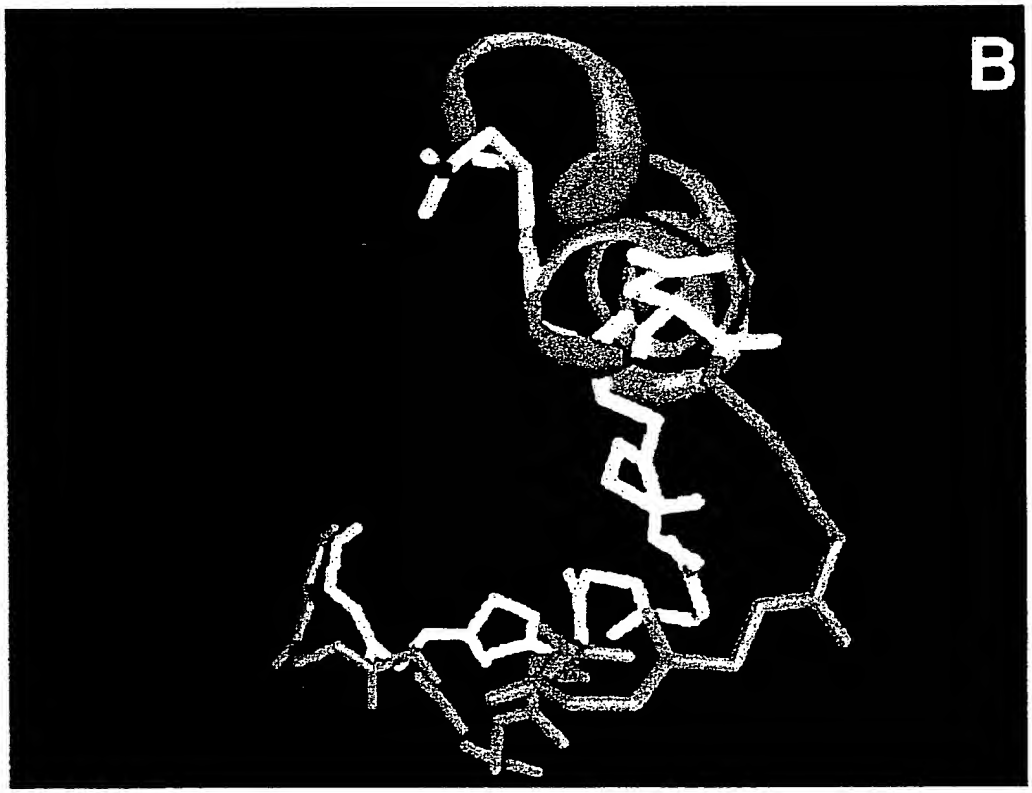


Figure 6 continued

INTERNATIONAL SEARCH REPORT

Intern al Application No
PCT/GB 00/01571

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/81 A61K38/55

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
BIOSIS, EPO-Internal, WPI Data, EMBASE, MEDLINE, PAJ, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	ALVAREZ-FERNANDEZ MARCIA ET AL: "Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 274, no. 27, 2 July 1999 (1999-07-02), pages 19195-19203, XP002144223 ISSN: 0021-9258 the whole document	1-10

-/--

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 August 2000	17/08/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Cervigni, S

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/GB 00/01571

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHEN JINQ-MAY ET AL: "Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases." FEBS LETTERS, vol. 441, no. 3, 28 December 1998 (1998-12-28), pages 361-365, XP002144224 ISSN: 0014-5793 cited in the application page 365, column 2, last paragraph	
A	CHEN J -M ET AL: "CLONING, ISOLATION, AND CHARACTERIZATION OF MAMMALIAN LEGUMAIN, AN ASPARAGINYL ENDOPEPTIDES" JOURNAL OF BIOLOGICAL CHEMISTRY,US,AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, vol. 272, no. 12, 21 March 1997 (1997-03-21), pages 8090-8098, XP002065429 ISSN: 0021-9258 cited in the application	
A	VITO TURK ET AL: "THE CYSTATINS: PROTEIN INHIBITORS OF CYSTEINE PROTEINASES" FEBS LETTERS,NL,ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, vol. 285, no. 2, 22 July 1991 (1991-07-22), pages 213-219, XP000215713 ISSN: 0014-5793	

1

HPS Trailer Page
for
WEST

UserID: stic_fpas

Printer: cm1_9e12_gbefptr

Summary

Document	Pages	Printed	Missed	Copies
WO000064945	45	45	0	1
Total (1)	45	45	0	-